Commentary on D.W. Massaro Speech-perception by ear and eye. A paradigm for psychological inquiry

de De Gelder, B.; Vroomen, J.H.M.

Published in:
Behavioral and Brain Sciences

Publication date:
1989

Commentary on D.W. Massaro, "Speech perception by ear and by eye:
paradigm for psychological inquiry " ”Behavioral and Brain Sciences"• (1989)

Models in the mind, modules on the lips.

Beatrice de Gelder and Jean Vroomen
Departments of Psychology and Philosophy,
Tilburg University, P.O. 90153,5000 LE Tilburg,
The Netherlands.

As a consequence of studies of both sign language and lip-reading it is now widely agreed that a comprehensive theory of speech perception (and production) must capture the common properties of the different modalities in which speech can be perceived. Generally speaking, evidence from sign-language and from lip -reading has been welcomed by modularist speech theorists. In contrast, Massaro's contribution builds on more than a decade of his research on speech perception in the auditory and visual modality and arrives at a refutation of modularism. Modularists welcome evidence from lipreading because they regard it as support for the view that speech perception draws upon a representational repertoire which is both language specific and a-modal or abstract. Massaro believes that the bi-modality of speech perception offers a spectacular insight into the general way in which the organism operates in its information processing, e.g. by integrating multiple sources of information. The existence of this general process is the core of his refutation of modularism. In this commentary we make two remarks about the relevance of lip-reading research to these issues. We believe that the issue of information integration is orthogonal to the matter of modularity of speech perception. What matters is domain-specificity of the representations over which integration processes operate. Next, we consider the relevance of language-specific deficits in the visual and auditory modality observed in developmental dyslexia.

1. Modularity is about representations.

Massaro claims that information integration is the basic principle of information processing in all domains of cognitive functioning. What aspects of the claims of modularism as epitomized in Fodor (1983) does this refute? The core of modularism is the

view that all cognitive processes are inferential but that some of them are subserved each by a special purpose computational system having a restricted databases. So both modularism and Massaro agree that information processing is hypothesis testing and decision making. Going beyond this, modularism proposes a cognitive typology of decisions. Some decision processes are local because they are based on limited knowledge (e.g. they are domain-specific and encapsulated). Massaro claims that a general model of decision making is enough, that one does not need to go through the trouble of designing a (modular) model of the internal knowledge environment of the organism to understand how decisions on inputs are reached. This is an empirical issue. The nexus of the disagreement is categorical vs. continuous perception. In Massaro's view categorical identification results from an integrative decision about a previous continuous evaluation process taking place separately in various dimensions. Modularism claims that evaluation is itself a decision process located in the language module and resulting in phonological representations. This view leaves room for an ulterior post-modular decisions about modular phonological representations. How does evidence for modality-specific coding square with either view? The answer is complicated by the fact that the notion of dimension and that of modality are orthogonal. Massaro uses the multi-modality of speech input as privileged example to show that iiformation processing consists of integration of input coming from various dimensions. This way he indicates that he treats modalities as dimensions and implies that phonological category decisions require integration of bi-modal information. In normal circumstances both the visual and the auditory modality are present. Yet, except for the cases where auditory information is deteriorated, it is perfectly possible to identify speech in the absence of information from the visual modality. Perception in the auditory modality still results from integration of multiple features. Perception of sign language in congenitally deaf subjects offers a similar example of multiple dimensions of a stimulus within one modality. The real issue is categorical perception within one and the same modality.

2. Developmental dyslexia: The discovery of a series of domain-specific impairments in cognitive functioning

Detailed studies of acquired dyslexia support the view that we are dealing with a specific impairment in the domain of written language skills (Shallice, 1988) which leaves intact decision processes in other domains of linguistic and non-linguistic information processing. Such disorders support modularism because as noted above, modularism is a

thesis about the nature of representational resources required in decision processes. There is a similar convergence towards the language-specificity of the deficits observed in developmental reading disorders. Obviously, developmental dyslexia manifests itself on the occasion of reading instruction. What causes these reading acquisition problems? We have explored the idea that they might be related to spoken language deficits. If so, adding an exploration of speech perception in the visual modality might get us either an independent check of the speech processing abilities of young dyslexics or, alternatively, reveal the existence of bi-modal integration difficulties. We could use a stimulus tape made by Massaro (Massaro 1983) to examine this issue. Our results (de Gelder and Vroomen, 1988) show that young dyslexics have less robust speech categories when their performance is compared with that of control groups (Both reading age and chronological age). This finding confirms the results by Werker and Tees (1986). Our subjects also lag behind in processing speech information in the visual modality. These modality based deficits show a significant correlation which suggests that there is an early integration of visual and auditory information. Does the McGurk-illusion also reveal early integration or on the contrary does it show late categorization within each modality? When presented with a auditory /ba/ and a visual /da/, the subject reports hearing what is a fusion of both, a /da/. On presentation of an auditory /da/ and a visual /ba/, what the subject reports is a blend, /bda/. Massaro's model(which does not give the full details of the representations activated at the various stages in processing) explains the former but might have difficulty with the latter case. He treats the perception of /bda/ as the identification of a single integrated speech event which is no exception to the general integration formula. The /b/ in /bda/ is indeed an integrated percept because subjects never report /mda/ or /pda/. There is, however, no visual influence in the /da/ part of /bda/. In other words, there is conflict plus integration. The extent to which there can be conflict between modalities would appear to be closely related to the robustness of the representations in each modality. Indirect evidence about robustness of modality specific coding has been obtained by using a serial recall paradigm (Massaro 1987:50). Adapting the paradigm of the auditory recency effect in serial recall Campbell and Dodd (1980) found that lip-read lists show recency just like auditory presented lists. This finding suggests that there is a common language source for both modalities. In a recent study using a large population of subjects we find that a visual suffix has no influence upon the recency of an auditory presentation. More surprisingly, an auditory suffix with no visual articulation does not affect recency in a visually presented list (de Gelder and Vroomen, 1989). This result

suggests that besides integration of visual and auditory information, there is still room for modality-specific coding.

ACKNOWLEDGEMENTS

Research reported was supported by a research grant to the first author. We thank Jose Morais and Paul Bertelson for comments on the manuscript.

REFERENCES

Bertelson, P. and de Gelder, B. Learning about reading from illiterates. In: A. Galaburda (Ed.) ”From neurons to reading.• Cambridge (Mass.) : MIT/Bradford Books.

Campbell, R., and Dodd, B. (1980) Hearing by eye. ”Quarterly Journal of Experimental Psychology, 32, 509-515.

de Gelder, B. and Vroomen, J. Cross-modal speech perception in young dyslexics. Paper presented at the ”6th Language and speech conference•,Sydney, aug. 1988.

de Gelder, B. and Vroomen, J. Seeing lips and hearing voices separately. ”•Workshop on Lipreading, Oxford”•, 28 april 1989.

Massaro, D.W. (1987) ”Speech perception by Ear and Eye : a Paradigm for Psychological Inquiry. Hillsdale: Lawrence Erlbaum.

Werker, J.F., and Tees, R.C. (1986). Speech perception in severely disabled and average reading children. ”Canadian Journal of Psychology, 41,• 48-61.